Run Time Complexity

- In typical application the total run time of a genetic algorithm is determined by the number of fitness function evaluations.
- Run time of selection algorithm and variation operators can be ignored.
- Number of fitness function evaluations is proportional to the number of generations run and the population size:

Fitness Fct. Evals = # Generations . Population Size

Convergence speed

- Rate at which a population converges is determined by the selection pressure:
  - high selection pressure: fast convergence
  - low selection pressure: slow convergence
- Size of population determines quality of solution found:
  - large population size: more reliable convergence
  - small population size: less reliable convergence
- Minimal fitness function evaluations: trade-off between selection pressure and population size

Key questions

1. How long does a GA - with a certain selection pressure - run before it converges ?
2. What is the minimal population size to ensure reliable convergence ?

   → ie. problem dependent, but:

- we can build analytical models for simple problems,
- use this as an approximation for real, complex problems,
- gives insight in and guidance for designing high performant GAs.

Models

1. First, we will build analytical models for the convergence behavior, assuming large enough (∞) populations,
2. Second, we will build analytical models for the minimal required population size,
3. Third, we will test the models on a real, complex problem (map labeling).

Selection Intensity

- To quantify the speed of convergence caused by the selection pressure we need a measure,
- The field of Quantitative Genetics already works with such a measure: the selection intensity I.

Quantitative Genetics

- Quantitative genetics studies the inheritance of those differences between individuals that are quantitative rather than qualitative.
- Quantitative differences have a continuous nature such as the height or the weight of the human body, whereas qualitative variation is measured in discrete units or categories such as eye color or blood type.
- To characterize the evolution of the quantitative differences the following concepts are defined.
- The **selection progress** or **response to selection** $R(t)$ is defined as the difference between the mean fitness of the population at generation $t+1$ and the mean fitness of the population at generation t.
- The **selection differential** $S(t)$ is the difference between the

mean fitness of the parent population at generation t and the population mean fitness at generation t. The parent population is the pool of individuals remaining after selection has been applied but before the oofspring has been generated by the variation operators:

$$S(t) = \overline{f(t^s)} - \overline{f(t)}.$$

- Assuming that the population fitness is normally distributed $N(\overline{f}, \sigma^2)$ we can scale the selection differential by the standard deviation $\sigma(t)$.
- This scaled selection differential is called the **selection intensity** $I(t)$. This is a dimensionless number since the standard deviation has the same units as the selection response:

$$I(t) = \frac{S(t)}{\sigma(t)} = \frac{\overline{f(t^s)} - \overline{f(t)}}{\sigma(t)}.$$

- Standardizing the normal distribution ($\overline{f} = 0, \sigma = 1$) shows that the selection intensity I is simply the expected average fitness of the population after applying the selection scheme to a population with standardized normal distributed fitness ($N(0,1)$).
- The relation between the response to selection R and the selection differential S is given by the heritability h^2:

$$R(t) = h^2 S(t),$$

or

$$R(t) = h^2 \sigma(t) I(t).$$

Proportionate selection

- Call $P_i(t)$ the proportion of occurrences of individual i in the population at generation t,
- Individual i has fitness f_i, and the mean fitness of the population at generation t is $\overline{f(t)}$,
- Call $P_i(t^s)$ the proportion of individual i in the parent pool after applying proportionate selection:

$$P_i(t^s) = P_i(t)\frac{f_i}{\overline{f(t)}}.$$

- The selection differential of proportionate selection is:

$$\begin{aligned} S(t) &= \overline{f(t^s)} - \overline{f(t)} \\ &= \sum_{i=1}^{N} P_i(t^s) f_i - \overline{f(t)} \end{aligned}$$

$$\begin{aligned} &= \sum_{i=1}^{N} P_i(t)\frac{f_i^2}{\overline{f(t)}} - \overline{f(t)} \\ &= \frac{1}{\overline{f(t)}}(\overline{f^2(t)} - (\overline{f(t)})^2) \\ &= \frac{\sigma^2(t)}{\overline{f(t)}} \end{aligned}$$

(i.e. Fisher's Fundamental Theorem of Natural Selection)

- note: $\sum_{i=1}^{N} P_i(t) f_i = \overline{f(t)}$; $\quad \sum_{i=1}^{N} P_i(t) f_i^2 = \overline{f^2(t)}$

$$\begin{aligned} \sigma^2 &= \sum_{i=1}^{N} \frac{(X_i - \overline{X})^2}{N} \\ &= \frac{\sum_{i=1}^{N} X_i^2}{N} - 2\overline{X}\frac{\sum_{i=1}^{N} X_i}{N} + \frac{N\overline{X}^2}{N} \end{aligned}$$

$$= \overline{X^2} - \overline{X}^2$$

- The selection intensity $I(t) = \frac{S(t)}{\sigma(t)}$ of proportionate selection is thus equal to the ratio of the standard deviation of the fitness and the population mean fitness:

$$\boxed{I(t) = \frac{\sigma(t)}{\overline{f(t)}}}$$

- Observations:
  1. The selection intensity of proportionate selection reduces if the fitness variance between the individuals in the population reduces and/or if the mean fitness increases. Both typically happen at later generations when the population starts to loose its diversity. The selection pressure basically disappears.

i	i_1	i_2	i_1	i_1	i_2	i_3	i_4	i_2	i_3	i_1
f_i	105	101	105	105	101	93	91	101	93	105

Mean fitness $\overline{f(t)} = 100$, proportions: $P_i(t)$, expected proportions after selection: $P_i(t^s)$, most likely number of copies: $n(t^s)$.

i	$P_i(t)$	$P_i(t^s)$	$n(t^s)$
i_1	0.4	0.420	4
i_2	0.3	0.303	3
i_3	0.2	0.186	2
i_4	0.1	0.091	1

$\Rightarrow$ most likely no change in number of copies !

Selection intensity $I(t) = \frac{\sigma(t)}{\overline{f(t)}} = \frac{5.31}{100} = 0.0531$

Expected mean fitness after selection (i.e. the parent pool):
$\overline{f(t^s)} = \overline{f(t)} + \sigma(t)I(t) = 100 + (5.31)(0.0531) = 100.28$.

2. The selection intensity of proportionate selection changes when the fitness values are transformed with a constant term, for instance when changing temperature values from Celsius to Fahrenheit (F = 1.8 C + 32).

Truncation Selection

- Selection differential $S(t) = \overline{f(t^s)} - \overline{f(t)}$,
- Truncating a normal distribution at the top τ% results in a mean fitness increase that is proportional to the standard deviation:

$$\overline{f(t^s)} - \overline{f(t)} = C(\tau)\sigma,$$

- By the definition of the selection intensity, $S(t) = I\sigma(t)$, it follows that $I(\tau) = C(\tau)$,
- For a given trunction threshold τ the selection intensity is a constant equal to the mean value of the right part of a standard normal distribution ($\overline{f(t)} = 0, \sigma(t) = 1$), truncated at the top τ%. Values can be calculated or looked up in tables:

τ	1%	10%	20%	40%	50%	80%
$I(\tau)$	2.66	1.76	1.2	0.97	0.8	0.34

- Contrary to proportionate selection, the selection pressure with truncation selection remains constant and is independent of the population mean fitness and variance (this is true for all ranked-based selection methods).

Tournament Selection

- Tournament selection with tournament size s: pick s solutions at random from the population, and select the solution with the best fitness.
- The selection intensity $I(s) = (\overline{f(t^s)} - \overline{f(t)})/\sigma(t)$ is equal to the expected value of the best ranked individual of a sample from s individuals taken from the standard normal distribution ($\overline{f(t)} = 0$ and $\sigma(t) = 1$).
- How to compute this ?

  $\rightarrow$ order statistics

Order Statistics

- Order statistics describes the statistical properties of a set of random variables that are ordered (or ranked) according to their value.
- Assume we take a random sample of size s of a population with a certain distribution probability, and we sort the sample in increasing order of magnitude:

$$x_{1:s} \leq x_{2:s} \leq \ldots \leq x_{s-1:s} \leq x_{s:s}$$

- The i^{th} order statistic is the random variable $X_{i:s}$ that represents the distribution of the corresponding value $x_{i:s}$.
- The probability density function $p_{i:s}(x)$ of the i^{th} order statistic $X_{i:s}$ gives the probability that the i^{th} ranked individual of a sample of size s will have a value equal to x.

- We need to compute this for the standard normal distribution,
- The probability density function is

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

and the cumulative distribution is

$$\Phi(x) = \int_{-\infty}^{x} \phi(x)dx.$$

- The probability density function $p_{i:n}(x)$ is given by:

$$p_{i:s}(x) = s \binom{s-1}{i-1} \Phi(x)^{i-1} \; (1-\Phi(x))^{s-i} \; \phi(x),$$

- The expected value $u_{i:s}$ of the i^{th} order statistic $X_{i:s}$ is:

$$u_{i:s} \quad = \quad \int_{-\infty}^{+\infty} x \; p_{i:s}(x) \; dx$$

$$= \quad s\binom{s-1}{i-1} \int_{-\infty}^{+\infty} x\phi(x)\Phi(x)^{i-1}(1-\Phi(x))^{s-i}dx$$

Tournament Selection

- The selection intensity $I(s)$ is equal to the expected value of the best ranked individual of a sample of s individuals taken from the standard normal distribution:

$$\boxed{I(s) = u_{s:s}}$$

s	2	3	4	5	6	7
$I(s)$	0.56 $(= \frac{1}{\sqrt{\pi}})$	0.85	1.03	1.16	1.27	1.35

- For a given tournament size the selection intensity is constant.

Binomial/Normal Distributed Fitness Function

- equal and additive genic fitness contributions
- uniform convergence at all loci
  $\Rightarrow$ lower bound on convergence complexity
- recombination makes no change to population mean fitness
  $\Rightarrow$ simple, yet accurate convergence models
- assume highly disruptive crossover: uniform crossover
- bit counting function $\Rightarrow$ fitness binomial distributed

Counting Ones fitness function

- Counting Ones, 'fruit fly' of GA theory

$$X = x_1 \dots x_\ell, \quad x_i \in \{0,1\}$$

$$CO(X) = \sum_{i=1}^{\ell} x_i$$

- Probability having 1 at a certain locus: $p(t)$
- Mean fitness at generation $t : \bar{f}(t) = l.p(t)$
- Variance at gen. $t : \sigma_p^2(t) = l.p(t)(1 - p(t))$

Proportionate Selection: Counting Ones

- mean fitness increase: $\overline{f(t+1)} - \overline{f(t)} = \sigma(t)I(t) = \frac{\sigma^2(t)}{\overline{f(t)}}$
- proportion of optimal alleles p(t)

$$p(t+1) - p(t) = \frac{1}{l}(1 - p(t))$$

$$\frac{dp(t)}{dt} \approx \frac{1}{l}(1 - p(t))$$

- convergence model ($p(0) = 0.5$)

$$\boxed{p(t) = 1 - 0.5e^{-t/l}}$$

- convergence speed: $p(t_{conv}) = 1 - 1/(2\ell)$

$$\boxed{t_{conv} = \ell \ln(\ell)}$$

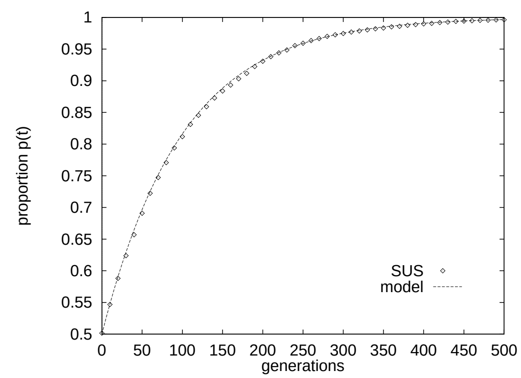

Figure 1: Convergence model and experimental results ($\ell = 100, N = 200$) for the Bit Counting problem using proportionate selection (Stochastic Universal Sampling) and uniform crossover.

Truncation Selection: counting ones

- mean fitness increase

$$\overline{f(t+1)} - \overline{f(t)} = \sigma(t) I(\tau)$$

- proportion of optimal alleles p(t)

$$p(t+1) - p(t) = \frac{I(\tau)}{\sqrt{l}} \sqrt{p(t)(1-p(t))}$$

$$\frac{dp(t)}{dt} \approx \frac{I(\tau)}{\sqrt{l}} \sqrt{p(t)(1-p(t))}$$

- convergence model ($p(0) = 0.5$)

$$\boxed{p(t) = 0.5(1 + \sin(\tfrac{I(\tau)}{\sqrt{l}} t))}$$

- convergence speed ($p(t_{conv}) = 1$)

$$\boxed{t_{conv} = \tfrac{\pi}{2} \tfrac{\sqrt{l}}{I(\tau)}}$$

- Compare $O(\sqrt{l})$ convergence time for truncation selection versus $O(l \ln(l))$ for proportionate selection.

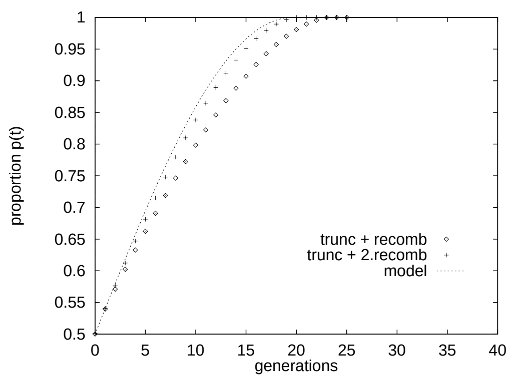

Figure 2: Convergence model and experimental results of the proportion of optimal bit values for the Bit Counting problem using truncation selection (pick 50 % best) and uniform crossover.

Tournament Selection: counting ones

- Mean fitness increase

$$\overline{f(t+1)} - \overline{f(t)} = \sigma(t) I(s),$$

- same equation as for truncation selection:
- convergence model ($p(0) = 0.5$)

$$\boxed{p(t) = 0.5(1 + \sin(\tfrac{I(s)}{\sqrt{l}} t))}$$

- convergence speed ($p(t_{conv}) = 1$)

$$\boxed{t_{conv} = \tfrac{\pi}{2} \tfrac{\sqrt{l}}{I(s)}}$$

- For instance, tournament size $s = 2 \Rightarrow I(2) = \frac{1}{\sqrt{\pi}}$
- convergence model ($p(0) = 0.5$)

$$\boxed{p(t) = 0.5(1 + \sin(\frac{t}{\sqrt{\pi l}}))}$$

- convergence speed ($p(t_{conv}) = 1$)

$$\boxed{t_{conv} = \frac{\pi}{2}\ \sqrt{\pi l}}$$

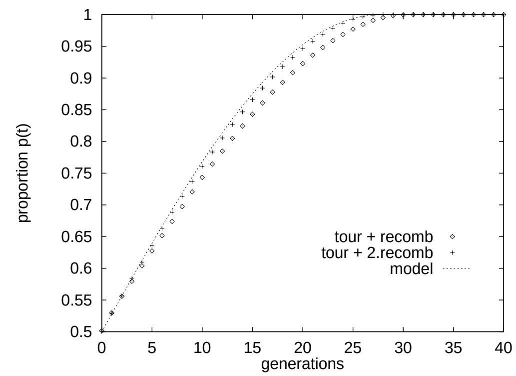

Figure 3: Convergence model and experimental results of the proportion of optimal bit values for the Bit Counting problem using tournament selection and uniform crossover.